Math 113 – 2020-Sep-14
3.2 Standard deviation

Given

Data set 1 $X = 26.7 \quad 29.6 \quad 80.5 \quad 4.8$

Data set 2 $X = 36.7 \quad 39.6 \quad 80.5 \quad 34.8$

Let's first find mean $\bar{x}$ (n (sample size))

Data set 1: $\bar{x} = (26.7+29.6+80.5+4.8)/4 = 141.6/4 = 35.4$

Data Set 2: $\bar{x} = (36.7+39.6+80.5+34.8)/4 = 141.6/4 = 35.4$

~~Even though~~ Data Sets 1 & 2 have the same mean but are different (subjectively) in that Data pts in set 2 are closer together

We need a way to describe mathematically how spread out our data is

Range (new)

Data set 1: range = max - min = 80.5 - 4.8 = 75.7

Data set 2: range = max - min = 39.6 - 30.5 = 9.1

Range is a quick & dirty way to measure "spread" of data.

Disadvantage of range: only uses 2 points, so doesn't take the whole data set into account.

New Topic Standard deviation s

Lets compute standard deviation s by hand.

Data set 1 $X = 26.7 \quad 29.6 \quad 80.5 \quad 4.8$

x	"deviations" $x-\bar{x}$	"square deviations" $(x-\bar{x})^2$
26.7	$26.7 - 35.4 = -8.7$	$(-8.7)^2 = 75.69$
29.6	$29.6 - 35.4 = -5.8$	$(-5.8)^2 = 33.64$
80.5	$80.5 - 35.4 = 45.1$	$(45.1)^2 = 2034.01$
4.8	$4.8 - 35.4 = -30.6$	$(-30.6)^2 = 936.36$

we already calculated $\bar{x} = 35.4$

sum of square deviations $= 3079.7$

sample variance $s^2 = \dfrac{\text{sum of square deviations}}{n-1} = \dfrac{3079.7}{3} = 1026.5667$

sample standard deviation $s = \sqrt{s^2} = \sqrt{1026.5667} = \boxed{32.04}$

Try Data Set 2: $x = 36.7\text{ ft}\ \ 39.6\text{ ft}\ \ 30.5\text{ ft}\ \ 34.8\text{ ft}$

(we expect s to be smaller than 32 because Data set 2 is closer together than dataset 1)

x	deviations $x-\bar{x}$		square deviations $(x-\bar{x})^2$	
36.7 ft	$36.7-35.4$	$=1.3$	$(1.3)^2$	$=1.69$
39.6 ft	$39.6-35.4$	$=4.2$	$(4.2)^2$	$=17.64$
30.5 ft	$30.5-35.4$	$=-4.9$	$(-4.9)^2$	$=24.01$
34.8 ft	$34.8-35.4$	$=-0.6$	$(-0.6)^2$	$=0.36$

secret 1st step: find $\bar{x}$ — we already calculated $\bar{x} = 35.4$

sum of square deviations $= 43.7$

sample variance $s^2 = \dfrac{\text{sum of square deviations}}{n-1} = \dfrac{43.7}{3} = 14.5667$

sample standard deviation $s = \sqrt{s^2} = \sqrt{14.5667} = \boxed{3.82}$

I'm going to round s to one more place than data

(This data has 1 decimal place, so for s I'll report 2 places)

Why square root at end? why not just use variance s^2?